

# IMAGE PROCESSING SYSTEM

## CROSS REFERENCE OF RELATED APPLICATION

This application is based on and claims priority under  
5 35 U.S.C. §119 with respect to Japanese Patent Application  
No. 2002-272523 filed on September 19, 2002, the entire content  
of which is incorporated herein by reference.

## BACKGROUND OF THE INVENTION

### 10 1. Field of the Invention

The present invention relates to a system for performing  
correlation process in order to synchronously reproduce video  
data (moving picture data) and still picture data.

### 2. Description of the Related Art

15 Services for distributing various types of contents to  
browsing clients, such as personal computers and mobile terminals,  
have been widely provided as specific area or wide area services.

There are various forms for the content distribution systems,  
and as the network communication technique has been developed,  
20 content data including video data are also distributed.

As the overview of the content distribution system, various  
content data are registered to a distribution server, and when  
a browsing client accesses the distribution server and selects  
and requests a desired content, the distribution server provides  
25 the requested content data for the browsing client.

In this process, in order to easily select the content, the video data distribution system correlates, with the video data, a content record (meta-data) that includes the title of the content, a keyword and category information to select  
5 the content and information for a content creator, so that the convenience for a user who employs the browsing client to watch the content is increased.

The types of video data to be distributed have been expanded to include a variety of materials, ranging from entertainments  
10 such as movies to educational references, lectures and visual presentations. Especially for lectures and visual presentations given with supporting reference materials, not only video pictures but also still pictures are provided for a browsing client, and the synchronous reproduction of these  
15 pictures is demanded, so that the information they provide can be fully employed to promote a better understanding of the content.

Further, for the video data for lectures that include many scenes, it is also requested that a specific scene in  
20 the video data where a reference is employed be searched for and reproduced. When such a search function can be performed by referring to the still picture for a representative scene in the video data, it is convenient because a desired scene can be easily and quickly found.

## SUMMARY OF THE INVENTION

The present invention has been made to solve the above-explained technical problems, and therefore, it is one object of the present invention to correlate video data with  
5 corresponding still picture data to enable synchronous reproduction of these data.

The present invention can be provided by various forms, such as an image processing system, an image processing method and a program that permits a computer to serve as the image  
10 processing apparatus or to perform the image processing method.

In any form, the correlation between video data and still picture data is performed to enable synchronous reproduction of these data.

According to an image processing system for this invention,  
15 a video display unit reproduces and displays video data on a screen, and an image display unit reproduces and displays still picture data on the screen. When, during the reproduction of the video data, a designation unit accepts an instruction from a user to designate a still picture displayed on the screen,  
20 a correlation unit correlates the designated still picture data with the reproduction time position in the video data.

Therefore, the still picture data, such as reference picture data, can be correlated at a reproduction time position of the video data in accordance with the user's desire. It is  
25 preferable that the video picture and the still picture be

displayed on the same screen, so that the user can enter a designation instruction at an accurate time.

The image processing system of the invention can register, to a distribution server that provides video data and still  
5 picture data in accordance with a request from a browsing client, the video data and the still picture data that are correlated by the registered client. In this case, the corresponding function is provided for the registered client.

The correlation function may be written as a program.  
10 Further, in accordance with the request from the browsing client, the distribution server may provide the video data and the program for the browsing client, and the browsing client executes the program as the video data are reproduced, and may request the distribution server the still picture data that are  
15 correlated at the reproduction time position.

The individual functions of the system may be provided as special devices, however, these functions can be implemented when a computer executes the program of the invention.

The image processing method of the invention can be  
20 implemented by using the image processing system of the invention.

#### BRIEF DESCRIPTION OF THE DRAWINGS

These and other objects and advantages of this invention  
25 will become more fully apparent from the following detailed

description taken with the accompanying drawings in which:

Fig. 1 is a diagram showing a system configuration according to one embodiment of the present invention,

Figs. 2A and 2B are diagrams for explaining the relationship  
5 between video data and picture data according to the embodiment of the invention,

Fig. 3 is a diagram showing a screen provided by a browsing client according to the embodiment of the invention,

Fig. 4 is a diagram showing a screen provided by the browsing  
10 client according to the embodiment of the invention,

Fig. 5 is a diagram showing a screen provided by the browsing client according to the embodiment of the invention,

Fig. 6 is a diagram showing a screen provided by the browsing client according to the embodiment of the invention,

15 Fig. 7 is a diagram showing a screen provided by the browsing client according to the embodiment of the invention,

Fig. 8 is a diagram for explaining a distribution server according to the embodiment of the invention,

Fig. 9 is a diagram for explaining the content data  
20 arrangement/registration processing according to the embodiment of the invention,

Fig. 10 is a diagram showing a screen provided by a registered client according to the embodiment of the invention,

Figs. 11A to 11C are diagrams for explaining the display  
25 style for a reference picture according to the embodiment of

the invention,

Fig. 12 is a diagram for explaining the Web function of the browsing client according to the embodiment of the invention.

5        DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The preferred embodiment of the present invention will be specifically explained.

Fig. 1 is a diagram showing a content distribution system according to the present invention. The content distribution  
10 system includes a distribution server 1, browsing clients 2 and a registered client 3, all of which are connected via the Internet N. Among these components, the present invention is mainly applied for the registered client 3.

The distribution server 1, the browsing clients 2 and  
15 the registered client 3 execute a program according to the invention by using computer hardware so as to perform predetermined processes, and the browsing clients 2 are constituted by personal computers especially having a browsing function for browsing contents.

20        The distribution server 1 stores, as each data content, video data and a content record (meta-data) including still picture data such as slide picture data or reference picture data correlated with the video data, voice index data for a search and a content title. When a request is received from  
25 the browsing client 2, the distribution server 1 distributes

corresponding content data.

Fig. 2A is a diagram showing the correlation between video data 5 and slide picture data 6, and Fig. 2B is a diagram showing the correlation between video data 6 and reference picture data 7.

Either the slide picture data 6 or the reference picture data 7 may be attached to the video data 5, and in the explanation for this specification, both of or one of the slide picture data 6 and the reference picture data 7 may be written as picture data, except when either data is specifically designated.

In this embodiment, the reference picture data 7 is defined as still picture data that an operator correlates with an arbitrary time position in video data, while reproducing the video data. The same definition may be applied for the slide picture data 6.

The video data 5 to be registered are moving picture data obtained by converting MPEG data into data stream for distribution, while the slide picture data 6 (A to K) are still picture data that are extracted automatically or by an operator from the video data 5 through the pre-process for the arrangement registration. The slide picture data 6 indicate representative pictures for a scene having a specific time width in the video, and are correlated with the corresponding scenes in the video data.

The slide picture data 6 are employed mainly for a search

process wherein the user searches for a desired content, or for a desired scene in the content video.

The reference picture data (a to n) are still picture data that the operator correlates with arbitrary reproduction time positions of the video data 5, while reproducing the video data 5 and watching the obtained pictures during the pre-process of the arrangement registration using the registered client 3. Example reference picture data are those for reference materials that are employed for a video presentation or a lecture that is presented by the video data 5.

As will be described later, the reference picture data 7 are correlated so as to be maintained on the screen until the time position whereat the next reference data 7 should be displayed is reached. As a result, the reference picture data 7 are correlated with the scene having a specific time width in the video representing the reference, and for the browsing user, the reference picture data can be reproduced and displayed in synchronization with the corresponding video data.

As will be described later, upon the reception of a request from the browsing client 2 that employs a browser, the distribution server 1 distributes the content data, such as the video data 5 and the still picture data 6 and 7, to the browsing client 2, and the data are displayed by the browsing client 2.



Fig. 3 is a diagram showing a default screen for the list of contents that are displayed by the browsing client 2 that has accessed the distribution server 1 at a predetermined URL.

The content list screen includes a search interface 10 and  
5 content interfaces 20 for individual contents. When a browsing user enters an instruction to the search interface 10 by using keys or by pointing at, a desired content can be found from the registered contents. When the browsing user enters an instruction to the content interfaces 20 by pointing at, the  
10 contents of the selected content data can be displayed and browsed.

For the search, the search interface 10 employs the meta-data or the voice index data that are registered in the distribution server 1 in correlation with the individual content  
15 data. The search interface 10 includes: a drop-down window 11 to select and enter a category; a keyword input portion 12 to search for desired content data; a search button 13 to enter an instruction to request the distribution server 11 to perform the search process based on the category and the  
20 keyword and to transmit the search results; a keyword input unit 14 to search for desired content data using the voice index data; and a search button 15 to enter an instruction to request the distribution server 1 to perform the search process based on the keyword and to transmit the search results.

25 The voice index data registered in the distribution server

1 in correlation with the content data is voice waveform data included in the content. The distribution server 1 converts the keyword entered at the keyword input portion 14 into voice waveform data, and compares the waveform data to search for  
5 content that includes the keyword.

Each of the content interfaces 20 includes a content number 21, a content title 22, a content creator 23, a content copyright owner 24 and a content category 25 based on the meta-data, and also displays a time scale 26 and a slide picture 27 for  
10 content video. The function for presenting the time scale 26 interlocks as a plug 26a is moved by a user, and changes the slide picture data 6 (A to K) to be displayed as a slide picture 27. As a result, the slide picture corresponding to the reproduction time position of the video data can be  
15 reproduced.

Further, a start button 28 and a detail button 29 are provided for the content interface 20. When the user points at the start button 28, the content video data can be obtained from the distribution server 1, and can be reproduced and  
20 displayed on the screen of the browsing client. When the user points at the detail button 29, as will be described later, the slide picture data or reference picture data for the content can be collectively obtained from the distribution server 1, and can be displayed as a list on the screen of the browsing  
25 client 2.

To cope with the video data 5 and all the correlated still picture data that are provided for the browsing client 2, the browsing client 2 has also a function that, when the user points at the slide picture 27 or a reference picture 30 on the list, the reproduction of the video data is started for a corresponding scene, i.e., at the corresponding reproduction time position, and the obtained picture is displayed.

For the content for which the correlated reference picture data are stored, the keyword search based on the reference picture data is enabled, and a content interface 20 representing the search results is shown in Fig. 4.

This content interface 20 is substantially the same as that in Fig. 3. In order to present these results, another function is also provided for the browsing client 2, whereby characters included in the reference picture 30 are compared with the input keyword and the list of the pertinent reference pictures 30 is displayed, and whereby, when one of the reference pictures 30 is pointed at by a user, the synchronous reproduction of the content video data is started at the corresponding scene, and the obtained pictures are displayed.

In accordance with the user's instruction, the browsing client 2 accesses the distribution server 1, receives desired content data, and displays the content data on a content browsing screen shown in Fig. 5. The content browsing screen has: a video player section 35 that includes a video screen 33, which

is used to display a reproduced video image, and a video operation  
portion 34 where operation buttons, such as reproduction and  
stop buttons, are arranged; an image display section 37, which  
has an operation portion 36 where a frame feeding button is  
5 arranged and is used to reproduce and display still picture  
data; and a note display screen section 38, which is used to  
display a program description concerning the content data.

Therefore, when the browsing user points at desired contents  
data, the video image can be reproduced and displayed on the  
10 video screen 33, and since the correlation process that will  
be described later has been performed, the reference picture  
(or the slide picture) corresponding to the reproduction time  
position of the video data can be synchronously displayed in  
the image display section 37.

15 Further, when the detail button 29 is depressed, the  
browsing client 2 accesses the distribution server 1, receives  
the image data of the desired content collectively, and displays  
these received image data on a details display screen shown  
in Fig. 6 or 7.

20 In Fig. 6, the list of the slide picture data 6 is shown.  
The items for the meta-data of the content are displayed in  
a record data display portion 40, and all the slide pictures  
6 correlated with the content are displayed in a time series  
manner in a slide picture display portion 41.

25 In this case, the sizes of the slide pictures 6 displayed

in the slide picture display portion 41 are increased in accordance with the time length of the corresponding scenes.

As is shown in Fig. 2A, for example, the slide pictures 6 C and D for which the time lengths of the scenes are shorter than the others are displayed smaller than the other slide pictures in the slide picture display portion 41. For this size conversion, the information about the time length of the scene need only be added to each slide picture, and the distribution server 1 need only change the picture size upon the reception of the details display request.

The image size may be set in accordance with the importance level of the scene, e.g., in accordance with the change of the contents of the scene or with the intent of a creator.

Fig. 7 is a diagram showing a screen for displaying the list of the reference picture data 7. The items of the meta-data of the content are displayed in a record data display portion 42, and in a reference picture display portion 43, all the reference pictures 7 correlated with the content are displayed in a time series manner, and character strings 7a that are segments extracted from the reference pictures 7 are also displayed. These character strings 7a are extracted as needed by the character recognition process, and are compared with the keyword to search for the reference picture, as is described above.

When a user points at one of still slide or reference

pictures on the list shown in Fig. 6 or 7, the video player 35 is displayed on the screen, and the video data can be reproduced beginning at the corresponding time position, and be displayed on the screen.

5       An explanation will be given for the processing for registering the video data and the correlated image data.

As is shown in Fig. 8, the distribution server 1 includes a database 50, for storing and managing the records of the individual registered contents; and a multiple location sets 10 51, for storing the substances of the content data.

In the database 50, stored are: content records 52, including the meta-data of the content data, the location file path and the location URL; and numbers 53 for specifying location sets where the content data are stored. When the database 15 50 is examined, the content data can be arranged in the folder of the location set 51, and the content data requested by the browsing client 2 can be distributed.

Each of the location sets 51 is a storage area to store one or multiple sets of content data, and a group of the location 20 sets 51 constitutes a content data storage area for the distribution server 1.

In the example in Fig. 8, for each location set 51 provided are: a video data folder 54, for storing video data; an image data folder 55, for storing still picture data correlated with 25 the reproduction time position of the video data; and a voice

index folder 56, for storing voice index data. The data of each content are stored in the corresponding folders 54 to 56, so that the data for the same content are stored in the same location set 51.

5       The arrangement and registration of the content data to each location set 51 is performed by an operator using a content preparation tool 60 of the registered client 3.

      The content preparation tool 60 performs: a process for converting MPEG video data into video data streams for  
10   distribution; a process for registering the converted video data; a process for registering the video data 5 in correlation with the slide picture data 6; a process for registering the video data 5 in correlation with the reference picture data 7, which will be described later; and a process for registering  
15   the video data 5 in correlation with voice index data.

      As is shown in Fig. 9, the process for arranging and registering the content data (video data, image data and voice index data) is initiated by entering to the registered client 3 content data 65 to be registered (1), and by connecting the  
20   registered client 3 to the distribution server 1.

      Then, the registered client 3 examines the set record 62 in the database 50, and sets the location set number of the content data of an archive file, the file paths of the media data and the URLs of the media data (2). Next, the  
25   registered client 3 transmits the content data to the

distribution server 1 for the registration (3).

At this time, the meta-data entered by the operator are also set in the content record 52, and the content record 52 and the set record 62 are correlated with each other based  
5 on the location set number.

Further, as will be described later, the correlation data (a program in this embodiment), which the registered client 3 has prepared in the process for correlating the reference picture data and the video data, is also transmitted to and  
10 registered by the distribution server 1. The correlation data are distributed together with the video data to the browsing client 2.

The content data that are thus arranged and located are provided by the distribution server 1, while the database 50  
15 is examined in accordance with the request from the browsing client 2.

Specifically, when the browsing client 2 issues to the distribution server 1 a request together with information about the designated contents, the distribution server 1 obtains  
20 the location set number of the content data by examining the content record 52, examines the set record 62 based on the location set number to obtain the URLs of the video data and the picture data, and provides the URLs to the browsing client 2. The browsing client 2 accesses content data files based  
25 on the URLs provided by the distribution server 1, and obtains



and displays the video data and picture data.

An explanation will now be given for the processing performed by the registered client 3 for correlating the video data 5 with the reference picture data 7.

5        Fig. 10 is a diagram showing a correlation interface displayed on the screen of the registered client 3. This interface is used when, as a pre-process for registration, the operator correlates the reference picture data with an arbitrary time position of the video data.

10        Before the correlation process is initiated by the operator, the video data file and the reference picture file are read by the registered client 3, and the correlation interface is displayed.

      This interface screen includes: a video player section  
15    72, which has a video screen 70 to display a video image obtained by reproducing the video data 5, and a video operation portion 71 where operation buttons, such as a reproduction button and a stop button, are arranged; an image display section 75, which  
      has a picture screen 73 to display a reference picture obtained  
20    by reproducing the reference picture data 7 and an operation portion 74 where a frame feed button and a decision button are arranged; a list display portion 76, to display the thumbnails of the reference pictures 30 included in the reference picture  
      file that is read; and a script display screen 77 to display  
25    correlation data (written as a program) obtained through the

correlation operation.

When on the interface screen the operator points at the frame feed button of the operation portion 74, an arbitrary picture in the reference pictures 30 in the list display portion  
5 76 can be expanded and displayed on the picture screen 73, so that the operator can confirm the details of the picture.

Further, when the operator points at the operation button of the video operation section 71, the video data 5 can be reproduced and the obtained video images can be displayed on  
10 the video scene 70. When the operator points at the decision button in the operation portion 74 during the reproduction of the video data 5, the reference picture data displayed on the picture screen 73 can be correlated with the pertinent reproduction time position of the video data 5.

15 Therefore, the operator can perform the processing while confirming the video picture and the reference picture on the same screen, and can correlate desired reference pictures with desired time positions of the video.

The registered client 3 has a function for, based on the  
20 instructions entered by the operator, writing a program that represents the correlation. When the operator manipulate the operation buttons, arbitrary number of reference picture data sets 7 can be correlated with arbitrary reproduction time positions of the video data 5, and these correlation data are  
25 written.

As is shown in Figs. 11A to 11C, based on the instructions entered by the operator, the registered client 3 can add, to the reference picture 30, an arrow 78 and a note 79 that are separately prepared. The reproduction time positions and the display positions for these added images are also included in the correlation data, and these added picture data are registered together with the reference picture data to the distribution server 11. It should be noted that, in the order from Figs. 11A to 11C, the image of the arrow 78 and the note 79 are added to the reference picture 30 as the video data are reproduced.

As is described above, the correlation program prepared by the registered client 3 is registered to the distribution server 1, and is provided together with the video data for the browsing client 2. An explanation will be specifically given for the synchronous reproduction process performed by the browsing client 2 for the video data and the program.

Fig. 12 is a diagram showing a Web page displayed by the Web browser of the browsing client 2. In the Web page, not only the video player 35 shown in Fig. 5 but also the correlation program that provided by the distribution server 1 are included.

This program includes a program code 80 and multiple sync command data sets 81.

The program code 80 includes: a moving picture reproduction position acquisition program, for obtaining

reproduction time position data for the video data; a sync command data check program, for examining the sync command data 81; and a sync command execution program, for executing a command corresponding to the sync command data 81.

5       The sync command data 81 includes video data reproduction time data indicating a command execution time, and command data to be executed. That is, each sync command data set 81 is present for the respective reference picture data correlated with the video data, and includes: the reproduction time  
10 position data indicating the position of the reference picture data correlated with the video data; and a command for requesting the distribution server 1 predetermined reference picture data when the reproduction time position is reached.

Therefore, the program code 80 is executed synchronously  
15 with the reproduction of the video data by the video player 35, and the moving picture reproduction acquisition program obtains the reproduction time position of the video data. Then, the sync command data check program determines whether there is the sync command data 81 that corresponds to the obtained  
20 reproduction time position. When such sync command data 81 is present, the command execution program executes the command in the sync command data, and requests the distribution server 1 for the corresponding reference picture data. This processing is repeated during the reproduction of the video data. The  
25 reference picture data that is provided is immediately reproduced

by the browsing client 2, and is displayed on the image display section 37, and as is shown in Fig. 5, the video data and the reference picture data are reproduced and displayed synchronously based on the correlation data.

5        In this embodiment, the reference picture is so set to be displayed on the image display section 37 until the display time position for the next reference picture is reached. Therefore, in the correlation process, the operator need only designate, for the video data, the start time position whereat  
10    the reference picture data is displayed. However, the start and end times for displaying the reference picture data may be entered, and at the designated time interval, the reference pictures may be displayed.

Further, in this embodiment, a program form to be executed  
15    by the browsing client 2 is employed for the correlation data.

However, the correlation data may be provided as a correlation table, and the browsing client may store all the necessary reference picture data, and when a reproduction time position is reached, may reproduce and display the corresponding reference  
20    picture data.

As is described above, according to the invention, since the correlation process for the video data and the still picture data is performed by a user while the video data and the still picture data are reproduced and displayed. Therefore, through  
25    the easy and precise manipulation of a user, the still picture

data can be correlated with an arbitrary reproduction time position of the video data, and the video picture and the still picture can be synchronously displayed at a time desired by the user.

5

[Fig. 8]

1: distribution server

2: browsing client

5

50: database

51: location set

54: video data holder

10      video data for content 1  
         video data for content 2  
         video data for content 3

55: image data holder

15      image data for content 1  
         image data for content 2  
         image data for content 3

56: voice index folder

20      voice index data for content 1  
         voice index data for content 2  
         voice index data for content 3

52: record for content 1  
record for content 2  
record for content 3

5 53: location number

60(3): content preparation tool (registered client) ✓

62: record for location set 1  
10 record for location set 2  
record for location set 3

[Fig. 9]

65: content data

15 ① input

3: registered client

② refer, set

③ store

50: database

20 62: set record

content 1

set number

video file path



image file path

voice file path

video URL

image URL

5 52: content record

53: content 1

title

keyword

set number

10 51: location set

set 1

content 1

video data file

image data file

15 voice index file